A NEW APPROACH TO SAR INTERFEROGRAM GEOCODING

Olaf Hellwich and Heinrich Ebner
Chair for Photogrammetry and Remote Sensing
Technische Universität München, D-80290 Munich, Germany
E-mail: {olaf|ebn}@photo.verm.tu-muenchen.de
URL: http://www.photo.verm.tu-muenchen.de
Commission I, Working Group 1/2

KEY WORDS: SAR Interferometry, Geocoding, DEM

ABSTRACT

Several geocoding methods for SAR interferometry are compared. The discussion results in the proposal of a new geocoding model. It is based on a least squares adjustment combining interferometric phase, range, Doppler centroid frequency, flight path and control point data. A complete mathematical framework for the computation of object space coordinates without approximations is presented. This gives a way to an efficient implementation of the algorithm for geocoding the pixels of an interferogram.

1 INTRODUCTION

Geocoding of SAR data has first been conducted for single scenes. A frequently used approach is the range-Doppler approach (Curlander, 1982; Raggam, 1988; Meier, 1989) where object space coordinates are computed from observed ranges, Doppler centroid frequencies, i.e. Doppler frequencies which occur when the object points are in the center of the radar beam, and flight path data. As range and Doppler centroid frequency determine the location of an object point in two dimensions only, additional information about the earth’s surface is introduced. This is possible when the flight path is given in a coordinate system referring to the earth’s surface, e.g. in a cartesian geocentric coordinate system. Two approaches are to be distinguished: in the first, a standard reference body of the earth, usually an ellipsoid, is introduced (Roth et al., 1993). The resulting location of an object point is correct only when the point is located close to the surface of the reference ellipsoid. All other points are subject to distortions particularly mountainous areas. This is avoided in the second approach where in addition to the reference ellipsoid a digital elevation model (DEM) is used to describe the surface of the earth (Meier et al., 1993). This geocoding procedure for SAR data is very similar to the derivation of orthophotos, a standard photogrammetric procedure (Kraus, 1995).

For geocoding SAR interferograms there is no need for additional information such as a reference ellipsoid. In combination with range and Doppler centroid frequency which determine the location of object points in two dimensions the interferometric phase determines the third dimension.

Figure 1 explains the geometric principle of SAR interferometry. An object point \( p \) is imaged twice from sensor positions \( s_1 \) and \( s_2 \) separated by baseline \( B \). The measurements taken from each sensor position are range \( r \), Doppler centroid frequency \( f_r \), and the complex amplitude

\[
V = \sqrt{I} \exp \{j\phi\}
\]

consisting of magnitude \( \sqrt{I} \), where \( I \) is called intensity, and phase \( \phi \). Range is determined by time measurement with an accuracy limited to e.g. 10 m for satellite-borne sensors. The range measurement defines a sphere on which an object point to be geocoded is located.

The Doppler centroid frequency is approximately 0. This means that the radar has a pointing direction more or less perpendicular to the flight path. In general, the Doppler centroid frequency defines a hyperboloid on which the object point is located. When the Doppler centroid frequency is precisely 0 the object point is located on a plane perpendicular to the flight path. Figure 2 shows the interferometric point determination for this case under the assumption of parallel flight paths. The drawing plane is the Doppler centroid plane to which the flight paths continue perpendicular. The spheres defined by range measurements are shown by circles with \( s_1 \) as center.

The phase components \( \phi_1, \phi_2 \) of the complex amplitudes of both SAR scenes are used to determine the range difference between the object point and the two sensor positions. The phase-based measurement has a high precision as it is computed from the difference \( \phi \) of the phases \( \phi_1 \) and \( \phi_2 \) of both SAR scenes. But \( \phi \) is ambiguous, because it is limited to the interval \( \left\{ -\pi, \pi \right\} \). This ambiguity can be solved by integrating phase differences \( \Delta \phi \) between neighbouring pixels, a process usually called phase unwrapping, and adding an integration constant \( \phi_0 \) determined, for instance, from a single control point. The range difference defines a hyperboloid which intersects the drawing plane of Figure 2 as a hyperbola on which the object point is located. In a general sense the range difference determines the elevation angle under which the object point is observed. Putting all the geocoded object points together a DEM can be derived.

The new approach combines the three observations range, Doppler centroid frequency and interferometric phase which are obtained for each interferogram pixel with flight path data to compute object space coordinates. First, the flight path data is refined by a least squares adjustment; then, the object space coordinates are computed using a multi-dimensional Newton-Raphson algorithm. In Section 2 previous approaches to the geocoding of SAR interferograms and the derivation of DEM from SAR interferometry
into the object space coordinate system to derive a procedure: first, terrain heights are computed from range centroid frequency and interferometric phase. Coordinates directly from sensor positions, range, Doppler subdivided into two groups. The first group uses a two-step approaches of the second group compute object space and interferometric phase difference. Then, the heights are integrated starting at a control point. (Hagberg and Lilander, 1993) use a similar approach. Instead of local terrain height differences incidence angle differences are computed. The integrated local incidence angles are then converted to object point heights using planar trigonometry.

None of the publications mentioned in the previous paragraphs treats the second step, the geocoding of height information. In all cases, either the earth surface is considered to be planar, or the height computation is treated referring to a tangential reference plane to the earth surface. Therefore, at least in the satellite-borne case, the descriptions are rather meant to be a treatment of the principles of SAR interferometry than a description of interferogram geocoding.

A more complete description of interferogram geocoding is given in (Small et al., 1993, Small et al., 1994). Control points are used to refine image and baseline geometry. The parameters estimated to refine the image geometry are range offset and scale expressed as corrections of the near range boundary and the range pixel spacing of the SAR scene, as well as time offset and scale in the form of scene start time and azimuth pixel spacing. The baseline parameters are a cross-track and a nadir bias parameter and a cross-track drift parameter. In addition to these parameters a phase constant is estimated from control points which is used to compute absolute interferometric phase from unwrapped interferometric phase. After the refinement the height is computed in a way similar to the approaches mentioned above. Planar trigonometry is used to compute object point height above a tangential plane centered in the imaged area. Applying cosine law the angle between the refined baseline and the look direction is computed. Then the look direction is used to compute height above the reference plane. The object point height is corrected with respect to an approximation of the earth's curvature. In the second step, the intermediate DEM is geocoded. For this task two different methods are used. The first method is the same as the one mentioned above for terrain-corrected geocoding of single SAR scenes using a reference DEM. In this case the interferometric DEM can be validated using the reference DEM. For practical applications this method does not appear to be very useful, as it requires a DEM as input which generally is output of the geocoding procedure. According to the second method the look vector is computed and added to the sensor position vector to compute object point coordinates. Unfortunately, the procedure of computations is not described.

(Prati et al., 1990, Hartl and Xia, 1993, Xia, 1996) treat the dual-pass sensor configuration. Using planar trigonometry terrain height differences of neighbouring object points are computed from differences of interferometric phases of neighbouring pixels as a function of baseline components and flying height above ground. As an approximation, in (Hartl and Xia, 1993, Xia, 1996) the directions from the sensor positions to the object point are assumed to be parallel. The computed height differences of neighbouring object points are integrated starting at a control point. (Hagberg and Lilander, 1993) use a similar approach. Instead of local terrain height differences incidence angle differences are computed. The integrated local incidence angles are then converted to object point heights using planar trigonometry.

None of the publications mentioned in the previous paragraphs treats the second step, the geocoding of height information. In all cases, either the earth surface is considered to be planar, or the height computation is treated referring to a tangential reference plane to the earth surface. Therefore, at least in the satellite-borne case, the descriptions are rather meant to be a treatment of the principles of SAR interferometry than a description of interferogram geocoding.

A more complete description of interferogram geocoding is given in (Small et al., 1993, Small et al., 1994). Control points are used to refine image and baseline geometry. The parameters estimated to refine the image geometry are range offset and scale expressed as corrections of the near range boundary and the range pixel spacing of the SAR scene, as well as time offset and scale in the form of scene start time and azimuth pixel spacing. The baseline parameters are a cross-track and a nadir bias parameter and a cross-track drift parameter. In addition to these parameters a phase constant is estimated from control points which is used to compute absolute interferometric phase from unwrapped interferometric phase. After the refinement the height is computed in a way similar to the approaches mentioned above. Planar trigonometry is used to compute object point height above a tangential plane centered in the imaged area. Applying cosine law the angle between the refined baseline and the look direction is computed. Then the look direction is used to compute height above the reference plane. The object point height is corrected with respect to an approximation of the earth's curvature. In the second step, the intermediate DEM is geocoded. For this task two different methods are used. The first method is the same as the one mentioned above for terrain-corrected geocoding of single SAR scenes using a reference DEM. In this case the interferometric DEM can be validated using the reference DEM. For practical applications this method does not appear to be very useful, as it requires a DEM as input which generally is output of the geocoding procedure. According to the second method the look vector is computed and added to the sensor position vector to compute object point coordinates. Unfortunately, the procedure of computations is not described.

(Prati et al., 1990, Hartl and Xia, 1993, Xia, 1996) treat the dual-pass sensor configuration. Using planar trigonometry terrain height differences of neighbouring object points are computed from differences of interferometric phases of neighbouring pixels as a function of baseline components and flying height above ground. As an approximation, in (Hartl and Xia, 1993, Xia, 1996) the directions from the sensor positions to the object point are assumed to be parallel. The computed height differences of neighbouring object points are integrated starting at a control point. (Hagberg and Lilander, 1993) use a similar approach. Instead of local terrain height differences incidence angle differences are computed. The integrated local incidence angles are then converted to object point heights using planar trigonometry.

None of the publications mentioned in the previous paragraphs treats the second step, the geocoding of height information. In all cases, either the earth surface is considered to be planar, or the height computation is treated referring to a tangential reference plane to the earth surface. Therefore, at least in the satellite-borne case, the descriptions are rather meant to be a treatment of the principles of SAR interferometry than a description of interferogram geocoding.

A more complete description of interferogram geocoding is given in (Small et al., 1993, Small et al., 1994). Control points are used to refine image and baseline geometry. The parameters estimated to refine the image geometry are range offset and scale expressed as corrections of the near range boundary and the range pixel spacing of the SAR scene, as well as time offset and scale in the form of scene start time and azimuth pixel spacing. The baseline parameters are a cross-track and a nadir bias parameter and a cross-track drift parameter. In addition to these parameters a phase constant is estimated from control points which is used to compute absolute interferometric phase from unwrapped interferometric phase. After the refinement the height is computed in a way similar to the approaches mentioned above. Planar trigonometry is used to compute object point height above a tangential plane centered in the imaged area. Applying cosine law the angle between the refined baseline and the look direction is computed. Then the look direction is used to compute height above the reference plane. The object point height is corrected with respect to an approximation of the earth's curvature. In the second step, the intermediate DEM is geocoded. For this task two different methods are used. The first method is the same as the one mentioned above for terrain-corrected geocoding of single SAR scenes using a reference DEM. In this case the interferometric DEM can be validated using the reference DEM. For practical applications this method does not appear to be very useful, as it requires a DEM as input which generally is output of the geocoding procedure. According to the second method the look vector is computed and added to the sensor position vector to compute object point coordinates. Unfortunately, the procedure of computations is not described.
SAR scenes is that the height $h$ of the object point above a geodetic reference ellipsoid is added to the equitorial and polar radii of the ellipsoid used. This means that each object point is located on the surface of its individual ellipsoid. On the basis of precise flight path data absolute interferometric phase can be simulated for given object points and their ellipsoidal heights $h$. By computing object space coordinates for the pixels of the interferogram using range $r$, Doppler centroid frequency $f$, and a set of assumed values $h$ a look-up table of $h = \tilde{h}(\phi,\psi)$ is simulated. It provides values $h$ as a function of pixel row $i$ and column $j$ and absolute interferometric phase $A$. Absolute interferometric phase is computed from unwrapped interferometric phase by addition of a phase constant derived with the help of at least one control point. The final height of an interferogram pixel is found by entering the look-up table with a phase value. Finally, point locations can be determined from the equations for ellipsoid-geocoding of single SAR scenes.

(Madsen et al., 1993) proposes a one-step geocoding algorithm applied to single-pass SAR interferograms. The algorithm is based on three-dimensional vector calculus. It uses the approximation that the look directions from the sensor positions to the object point are parallel. With the help of an orthogonal basis object space coordinates are computed from range, Doppler and interferometric phase equation. As this approach is similar to the new approach proposed here it is not described in further detail. The new approach differs from (Madsen et al., 1993), as it does not assume approximately parallel look directions. Consequently, a closed-form description of object space coordinates cannot be derived, and an iterative procedure is chosen for the computations. In contrary to (Madsen et al., 1993) the new approach refines the baseline information using control points.

3 A NEW CONCEPT FOR INTERFEROMGRAM GEOCODING

In this Section a brief description of the new approach is given which does not rely on any conceptual approximations and strictly uses three-dimensional geometry. It was sketched previously in (Hellwich, 1997). Object space coordinates are computed on the basis of range, Doppler and interferometric phase equation. Sensor positions are derived from flight path data which are refined by a least squares adjustment using control points. As range, Doppler and interferometric phase equation cannot be transformed to explicit formulas for object space coordinates, the interferogram pixels are geocoded using a multidimensional Newton-Raphson algorithm. Input parameters are the three observations per pixel (range, Doppler centroid frequency and interferometric phase) and the refined flight paths.

Image geometry can be refined according to (Small et al., 1994), an approach discussed in Section 2. As the refinement of image geometry can be conducted for each scene individually, it can be kept separate from the kernel geocoding operations.

Then, the least squares method is used to refine the sensor positions which is equivalent to a refinement of the baseline. As the elevation angle from a sensor position to the object point is computed from range difference as a function of interferometric phase, the baseline component in look direction has to be known with superior accuracy, for instance, to a fraction of a wave length. For satellite-borne dual-pass interferograms this is not the case as long as original orbit data determined according to the recent state of the art are used. The orbits are determined by an epoch state vector and an orbit model including several influences among which the earth's gravity field is most important. While the orbit model guarantees a relative accuracy sufficient for interferometric purposes, the epoch state vector is subject to errors in the range of $dn$. This means that the shape of the orbit is known, but that position and scale are not known accurately enough. The latter can be refined by bias and drift parameters determined with the help of control points.

The observation equations of the least squares adjustment are

- the interferometric phase equation (dual-pass case)
  \[ \phi + \hat{\phi} = \frac{4\pi}{\lambda} \left( |\tilde{p} - \tilde{s}_1| - |\tilde{p} - \tilde{s}_2| \right) + \phi_c \]  
  (1)

where $\phi$ is the observed unwrapped interferometric phase as a function of the sensor positions $s_1, s_2$, the object point $p$, the phase constant $\phi_c$ to transform unwrapped interferometric phase to absolute interferometric phase, and the wave length $\lambda$. $\hat{\phi}$ is the residual, and the "hat" (\(^{\hat{}}\)) indicates adjusted variables. The vectorial parameters consist of three coordinates $x, y$ and $z$. Note that Equation 1 states that the sensor positions are unknowns and as such determined in the course of the adjustment.

- the range equation
  \[ r_1 + \hat{r}_1 = |\tilde{p} - \tilde{s}_1| \]
  (2)

where $r_1$ is the range between sensor position 1 and the object point.

- the Doppler equation
  \[ f_1 + \hat{f}_1 = \frac{(\tilde{p} - \tilde{s}_1) \cdot \hat{v}_1}{\lambda (|\tilde{p} - \tilde{s}_1|)} \]
  (3)

where $f_1$ is the Doppler centroid frequency and $v_1$ is the velocity of sensor 1.

- the equations of flight path parameters
  \[ s_1 + \hat{s}_1 = \tilde{s}_1 + \tilde{b}_i + \tilde{d}_i t \]
  \[ v_1 + \hat{v}_1 = \hat{v}_1 \]
  \[ s_2 + \hat{s}_2 = \tilde{s}_2 + \tilde{b}_i + \tilde{d}_i t \]
  (4)

where $b_i$ is the vector of bias parameters in $x, y$ and $z$-direction, and $d_i$ the vector of drift parameters of flight path $i$. $t$ is the time of measurement which is assumed to be correctly set or corrected due to a preceding image refinement. The sensor positions $s$ are only formally considered unknown and to be determined by the adjustment. Practically, the shape of the orbit is to be preserved which means that $\tilde{s}$ is only influenced by $\tilde{b}$ and $\tilde{d}$, and not by its residual $\hat{s}$. This is achieved by introducing $s$ with a high weight.

- the control point equation
  \[ p + \hat{v}_p = \tilde{p} \]
  (5)
Observation equations could also be introduced for bias $b$ and drift $d$ whose accuracies are available from orbit determination. This would allow to control the magnitudes of the adjusted bias and drift parameters, a means which is considered to be not absolutely necessary.

The adjusted unknowns $\hat{x}$ are computed according to the well-known equations of least squares adjustment (Mikhail, 1976):

$$\hat{x} = \bar{x} + \Delta x$$  \hspace{1cm} (6)

where $\bar{x}$ is the vector of approximations of the unknowns and $\Delta x$ is the vector of corrections to the approximations. The linearized observation equation system is

$$\hat{v} = A\Delta x - w$$  \hspace{1cm} (7)

where $\hat{v}$ is the vector of residuals, $A$ is the Jakobi matrix containing the derivatives of the observations with respect to the unknowns, and $w = \sigma - \sigma(\hat{x})$ is the vector of shortened observations, i.e. the difference between observations $\sigma$ and observations as a function of the approximations of the unknowns $\hat{x}$. The "circle" (⊙) indicates approximated variables. The corrections to the approximations of the unknowns $\Delta x$ are resulting from

$$\Delta x = Q_{\Delta x}A^TPw$$  \hspace{1cm} (8)

where $P$ is the weight matrix. The covariance matrix of the adjusted unknowns $Q_{\Delta x}$ is computed according to

$$Q_{\Delta x} = \delta_0^2(A^TPA)^{-1}$$  \hspace{1cm} (9)

where $\delta_0^2$ is the a posteriori reference variance associated with the weight unity:

$$\delta_0^2 = \frac{\hat{v}^TP\hat{v}}{r}$$  \hspace{1cm} (10)

where $r$ is the redundant.

The adjustment is conducted for the control points only. This means that in the adjustment the bias and drift parameters, the sensor positions and velocities as well as the object space coordinates of the control points, and the phase constant $\phi_e$ are estimated. Note that only bias and drift parameters and $\phi_e$ are used in the following processing step. Object points which are not control points are not included in the adjustment, as they are not determined redundantly and, therefore, do not contribute to the determination of bias, drift and $\phi_e$.

Once bias and drift parameters refining baseline geometry are computed, all interferogram pixels are geocoded. First, the adjusted sensor positions $\bar{x}_1$ and $\bar{x}_2$ are computed from (4). This means that the sensor positions are interpolated and refined. Then, for each pixel a system of three equations (1), (2) and (3) with the three unknown object point coordinates $x_p, y_p$ and $z_p$ has to be solved. As there is no closed-form solution for this problem, an iterative solution is attempted. Here a multi-dimensional Newton-Raphson algorithm is applied (Press et al., 1992). It uses the Jakobi matrix $A_p$, basically by computing

$$\hat{p} = \bar{p} + \Delta p$$  \hspace{1cm} (11)

and

$$\Delta p = A_p^{-1}w_p$$  \hspace{1cm} (12)

where

$$w_p = \left(\phi - \phi(\bar{x}), r_1 - r_1(\bar{x}), f_1 - f_1(\bar{x})\right)$$

Simulations have shown that the convergence properties of the algorithm are very good owing to the geometric properties of the problem. Figure 2 shows that the solution space does not contain any severe irregularities. For the first unknown object point any location on the correct side of the baseline can serve as approximation. In about four iterations the algorithm provides a sufficiently accurate solution. In an interferogram geocoded pixel by pixel, any object point to be processed after the first one will be located close to a previous point which will serve as an approximation. The simulations conducted so far have shown that under these circumstances a single iteration always results in sufficiently accurate coordinates. This means that the algorithm provides a solution which is computationally equivalent to a closed-form solution of the equation system.

The resulting object space coordinates usually have to be transformed, e.g. from a geocentric cartesian coordinate system to a map projection coordinate system. In this coordinate system a DEM can be derived from the object points.

Summarizing, the proposed geocoding method consists of five steps:

1. Image refinement
2. Baseline refinement by least squares adjustment
3. Computation of object space coordinates for each interferogram pixel by Newton-Raphson method
4. Transformation to map projection coordinates
5. DEM generation

### 4 DISCUSSION AND OUTLOOK

The main advantage of the new geocoding method is that it is based on a consistent mathematical framework. A consequence of this framework is an implementation which allows fast geocoding of the bulk of interferogram pixels and a high potential for the analysis of accuracy.

A disadvantage may be the requirement for comparatively accurate flight path information. The question whether previously developed methods have an implicit advantage regarding a lower demand for flight path accuracy still has to be investigated. The new method has an obvious disadvantage concerning the introduction of control points, as their covariance matrices have to be transformed (rotated) from e.g. a topocentric to a geocentric coordinate system. Previously developed geocoding methods do not require such an operation.

Future work still has to be conducted concerning

- simulations of the adjustment for various baseline configurations,
- analysis of accuracy properties,
- tests with real SAR interferograms and
- comparisons with other geocoding methods under well-controlled conditions.
The authors apologize for the incompleteness of this work, yet hope that the new concept favours discussions about more consistent approaches to the geocoding of SAR interferograms.

REFERENCES


